DOCUMENT RESUME

ED 416 633 EC 306 184

AUTHOR Thurlow, Martha; Hurley, Christine; Spicuzza, Richard; Erickson, Ronald

TITLE Resources: Students with Disabilities in National and Statewide Assessments. State Assessment Series: Minnesota Report 7.

INSTITUTION Minnesota State Dept. of Children, Families, and Learning, St. Paul.; National Center on Educational Outcomes, Minneapolis, MN.

SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.

PUB DATE 1996-08-00

NOTE 35p.

CONTRACT R279A50011

AVAILABLE FROM Publications Office, NCEO, 350 Elliott Hall, 75 East River Road, University of Minnesota, Minneapolis, MN 55455; phone: 612-626-1531; fax: 612-624-0879; World Wide Web: http://www.coled.umn.edu/nceo (document may be copied without charge; additional print copies $8).

PUB TYPE Reference Materials - Bibliographies (131)

EDRS PRICE MF01/PC02 Plus Postage.

DESCRIPTORS *Disabilities; *Educational Assessment; Educational Policy; Elementary Secondary Education; Legal Problems; *Psychometrics; Research Needs; Statistical Analysis; Student Attitudes; *Student Evaluation; Teacher Attitudes; Test Reliability; Test Validity; Testing Problems; Testing Programs

IDENTIFIERS *Testing Accommodations (Disabilities)

ABSTRACT

This bibliography provides abstracts of 28 recent (most were published since 1993) books, journal articles, and other documents concerned with testing accommodations for students with disabilities. The emphasis was on studies examining the effects of testing accommodations on the technical integrity of assessment measures. However, the bibliography notes that, currently, comprehensive empirical studies of the effects of testing accommodations are still noticeably absent from the literature, but that three federally supported projects are currently addressing issues related to the assessment of students with disabilities. The literature reviewed is organized into four sections: (1) empirical studies of testing accommodations; (2) legal considerations related to testing and accommodations; (3) teacher and student perceptions of classroom and testing adaptations and modifications; and (4) conceptual issues. Appended are a list of current project titles and recipient organizations. (Contains 49 references from a 1993 bibliography.) (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 416 633

1

STATE ASSESSMENT SERIES
Minnesota Report 7

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Resources: Students with Disabilities in National and Statewide Assessments

BEST COPY AVAILABLE

EC 306 184

Minnesota Children

Minnesota Department of Children, Families & Learning

STATE ASSESSMENT SERIES
Minnesota Report 7

Resources: Students with Disabilities in National and Statewide Assessments

Minnesota Assessment Project

Project Staff:
Constance Anderson • Cathy Wagner • Michael Tillman
Minnesota Department of Children, Families and Learning

Project Advisors:
Karon Hergenhahn • Bounlieng Phommasouvanh • Marilyn Sorensen • Elizabeth Watkins
Minnesota Department of Children, Families and Learning

Prepared By:
Martha Thurlow • Christine Hurley • Richard Spicuzza • Ron Erickson
University of Minnesota

August 1996

The Minnesota Assessment Project is a four-year, federally funded effort awarded to the Minnesota Department of Children, Families and Learning from the U.S. Department of Education, Office of Educational Research and Improvement. The project's goal is to promote and evaluate the participation of students with limited English proficiency and students with disabilities in Minnesota's Graduation Standards. Specifically, the project will examine ways in which students with limited English and students with disabilities can participate in the Basic Standards Tests of reading, mathematics, and written composition and in the performance-based assessments of the high standards in the Profile of Learning.

This project is supported, in part, by a grant to the Minnesota Department of Children, Families and Learning from the U.S. Department of Education, Office of Educational Research and Improvement (Grant #R279A50011). Opinions expressed herein do not necessarily reflect those of the U.S. Department of Education or Offices within it.

Permission is granted to copy this document without charge. Additional print copies may be purchased for $8.00 from:

NATIONAL CENTER ON EDUCATIONAL OUTCOMES
University of Minnesota • 350 Elliott Hall
75 East River Road • Minneapolis, MN 55455
612.624.8561 • Fax 612.624.0879
http://www.coled.umn.edu/NCEO

This document is available in alternative formats upon request.

Overview

In 1993, the National Center on Educational Outcomes (NCEO) published a comprehensive literature review on testing accommodations for students with disabilities (Thurlow, Ysseldyke, & Silverstein, 1993). The results of that review documented that (1) very little empirical research existed on testing accommodations, and (2) there was tremendous variability across states in terms of the degree to which they included students with disabilities in assessments or made accommodations for them. A review of the literature published since that report suggests that in some ways, little has changed with respect to empirical research on testing accommodations. Currently, comprehensive empirical studies of the effects of testing accommodations are still noticeably absent from the literature on assessment and students with disabilities.

In 1995, the U.S. Office of Special Education Programs funded three projects to examine issues related to assessment for students with disabilities. Similarly, the U.S. Department of Education's Office of Educational Research and Improvement (OERI) funded eight states, including Minnesota, and one consortium of 22 states, to improve their state assessments through alignment of assessments with standards, and increased inclusion of students with disabilities and limited English proficiency in their assessments (Erickson, Thurlow, & Ysseldyke, 1996). Most of these projects are addressing testing accommodations. Appendix A provides a list of the project titles and recipient organizations.

The purpose of this annotated bibliography is to provide an up-to-date list of the literature on testing accommodations for students with disabilities with an emphasis on studies examining the effects of testing accommodations on the technical integrity of assessment measures. The literature is organized into four sections: (1) empirical studies of testing accommodations (pages 3–6), (2) legal considerations related to testing and accommodations (pages 6–8), (3) teacher and student perceptions of classroom and testing adaptations and modifications (pages 8–11), and (4) conceptual issues related to testing and accommodations (pages 11–18).

Sources of information for this literature review included books, journal articles, agency reports, personal communications with researchers involved in similar efforts in the field, documents published by research centers (e.g., North Central Regional Educational Laboratory [NCREL]) and testing companies (e.g., Educational Testing Service [ETS]), as well as papers presented at national conferences. The purpose of the searches was to locate empirical studies examining the effects of testing accommodations on the technical adequacy of measures. The current search maintained as criteria for inclusion either publication in or after 1993 (given the 1993 Thurlow et al. document) or if published prior to 1993, the source had not been included

in the 1993 review document. A list of references included in Thurlow et al. (1993) is included in Appendix B.

Searches used computer databases, including ERIC and PsycLit, the World Wide Web (using Alta Vista and Yahoo search engines), as well as a search of the Outcomes-Related Bank of Informational Text (ORBIT), a computerized literature database maintained by NCEO at the University of Minnesota. The following keywords, listed here in alphabetical order, were used in various combinations to conduct database searches: accommodations, adaptations, assessment, competency tests, disabilities, effectiveness, empirical studies, graduation standards, high stakes assessment, measurement, modifications, psychometric properties/qualities, reliability, special education, standards, technical adequacy, test(s, ing), and validity. Finally, the Social Citations Index was also reviewed for the purpose of finding current research that had referenced the NCEO report or any of the seminal articles from that review.

Dalton, B., Morocco, C.C., Tivnan, T., & Rawson, P. (1994). Effect of format on learning disabled and non-learning disabled students' performance on a hands-on science assessment. *International Journal of Educational Research, 21* (3), 229-316.

Prior research has shown that students with learning disabilities (LD) often struggle with multiple choice tests and the skills that they require. In this study, the effects of two alternative assessments – a constructed diagram test and a written questionnaire – were compared for 172 fourth grade students with (N = 33) and without (N = 139) learning disabilities from six urban and two suburban classrooms. Results show that students' outcomes were a function of learner status (LD, low, average, and high achieving) and level of science knowledge after instruction. Students with learning disabilities, and low and average achieving students obtained higher scores on the constructed diagram test than on the questionnaire after controlling for domain-specific knowledge. High achieving students performed comparably on the two measures. The majority of students (88%) reported that they liked the diagram test better, stating that it was fun and easier than the questionnaire. Possible explanations for differential performance include the possibility that the two tests measured different aspects of achievement or perhaps that the diagram test scaffolds student performance. The authors concluded that it is important to use a variety of sources and measures to assess student performance.

Dunlap, G., Foster-Johnson, L., Clarke, S., Kern, L., & Childs, K.E. (1995). Modifying activities to produce functional outcomes: Effects on the problem behaviors of students with disabilities. *Journal of the Association for Persons with Severe Handicaps, 20* (4), 248-258.

The purpose of the experiment was to determine whether problem behaviors exhibited by three elementary students with disabilities (including autism, mental retardation, and emotional behavioral disorders) could be reduced and on-task behavior increased if students' curricular activities were modified according to their own interests. For each student, a particular instructional objective was held constant by the way in which the object was modified to make the task more interesting to the student. Using a reversal design, results showed that all three students reduced problem behaviors and increased on-task behaviors when their curricular tasks were modified according to their interests. In their discussion, the authors asserted that although the conceptual basis for the changes in student behavior are not fully understood at this time, the functional outcomes are what is important.

Mick, L.B. (1989). Measurement effects of modifications in minimum competency test formats for exceptional students. *Measurement and Evaluation in Counseling and Development, 22*, 31-36.

This investigation examined the effects of three format modifications to the IOX Basic Skill Test (Reading, Secondary level), on the achievement of 76 secondary students with learning disabilities (LD) and mild to moderate mental handicaps (MMMH). The modifications were:

- Moderately increased print size.
- Use of unjustified lines for right margins.
- Responses recorded on test booklets rather than answer sheets.

Using a repeated replication design, the unmodified and modified versions of the test were administered to each student. Results indicated that both students with LD and MMMH performed significantly better on the unmodified version of the test. One possible explanation put forth by the author was that secondary students had become test-wise after long-term exposure to standardized test formats (including answer sheets and justified margins). Two relevant limitations were a small sample size and having to produce two equivalent forms of a test.

Munger, G.F. & Loyd, B.H. (1991). Effect of speededness on test performance of handicapped and nonhandicapped examinees. *Journal of Educational Research, 85* (1), 53-57.

The purposes of this study were to investigate whether performance on the Iowa Tests of Basic Skills (ITBS) (Language Usage and Expression and Mathematics Concepts) was related to test speededness for students with and without disabilities and to determine whether reducing the amount of speededness had differential effects on the two groups. Two hundred twenty-two fifth graders with and without disabilities were administered the ITBS under timed and non-timed conditions. Results indicated that there were no differences in performance under timed conditions nor were students differentially affected when the amount of speededness was reduced. These findings appear to be consistent with other research reviewed by the authors in which the manipulation of timing produced no significant effects. Based on their results, the authors concluded that changes in testing time may not affect student performance; therefore, schools should consider including students with disabilities in standardized testing.

Perlman, C., Borger, J., Collins, C., Elenbogen, J., & Wood, J. (1996, April). *The effect of extended time limits on learning disabled students' scores on standardized reading tests*. Paper presented at the annual meeting of the National Council on Measurement in Education, New York, New York.

Eighty-five fourth (N=28) and eighth-grade (N=57) students with learning disabilities were given the Iowa Tests of Basic Skills (ITBS) in either a timed or untimed administration. Analysis of covariance results, using previous ITBS scores as the covariate, indicated that there were significant main effects for both timing and grade; students in the untimed condition as well as students in eighth grade scored significantly higher than students in the timed condition or fourth graders. Additional findings suggest that the post-test was more reliable when untimed, students in the untimed condition did not always use all of the allotted time, and older students were more likely to need extra time. Moreover, fourth graders in the untimed condition tended to score higher than the fourth graders in the timed condition, even though they both used about the same amount of time. This particular result caused the authors to speculate as to whether the critical variable is time, or if reduced stress and more positive expectations accompany a student knowing that he or she has unlimited time. However, the discussion of findings was tempered by methodological concerns that significantly limit the study.

Ragosta, M. & Wendler, C. (1992). *Eligibility issues and comparable time limits for disabled and nondisabled SAT examinees*. College Board Report No. 92-5. ERIC #ED349337.

As a follow up to research conducted by Willingham et al. (1988), the purposes of this study were to (1) establish empirically derived testing times for special administrations of the SAT for examinees with disabilities, and (2) develop eligibility guidelines for students taking special administrations. Using data from the 1986-87 and 1987-88 SAT test administration timing records (N = 17,632), the SAT history file, and a survey questionnaire, Ragosta and Wendler determined that in general, comparable numbers of students with disabilities and students without disabilities completed the exam when students with disabilities were given between time-and-a-half to double time for special test administrations. For students who were blind and required Braille or cassette tape administrations, two to three times the standard testing time was required for similar completion rates.

With respect to guidelines for special administrations, the authors put forth an eligibility hierarchy based on school practices (e.g., a continuum whereby students with current IEPs would clearly qualify for accommodations while students with disabilities who have no history

of needing special accommodations in school would not be likely to qualify) and used the hierarchy to examine students who had taken the SAT in 1986-87 or 1987-88. Results indicated that some types and levels of disability could be differentiated using these criteria (e.g., hearing and visual impairments), while others could not (e.g., learning disabilities). Additional results documented that a significant number of students who were eligible for accommodations took the standard SAT and that another substantial group of students who had no record of receiving accommodations in school took the accommodated test. The authors concluded with a discussion of possible policy changes, including using a timing policy for special administrations based on empirically derived time limits and changing to school-based criteria for eligibility decisions.

Legal Considerations Related to Testing and Accommodations

The primary source of information on legal issues pertaining to testing accommodations is Dr. S.E. Phillips, a professor at Michigan State University. Specializing in legal issues in assessment and psychometrics, Dr. Phillips also holds a law degree. She is the author of all four of the articles reviewed in this section. Understandably, there is considerable overlap across the articles. In each article Phillips reviews (in various levels of detail) the federal statutes and case laws related to the topic of educational testing and accommodations. She also discusses psychometric considerations related to testing accommodations. Although the following annotated bibliographies note the statutes and laws Phillips reviews, the focus is more on describing the unique attributes of each article.

Phillips, S.E. (1993). Testing condition accommodations for disabled students. *Education Law Reporter, 80*, 9-32.

Focusing primarily on diploma-sanction tests, Phillips examined relevant legal precedents and psychometric standards as reflected in case law and the professional literature to provide guidance to test administrators who make decisions about testing accommodations. She reviewed three federal legal precedents for the provision of accommodations for students with disabilities: Constitutional due process, IDEA, and ADA. She then reviewed state precedents, specifically examining two state decisions regarding testing accommodations: *Board of Education of Northport v. Ambach* and *Hawaii State Department of Education* (1990). Phillips discussed the purpose of diploma-sanction testing as well as the challenges associated with the classification of students as learning disabled. She then discussed psychometric considerations, focusing specifically on validity, valid and invalid accommodations, and making

determinations about accommodations. Finally, she discussed the challenge of balancing individual rights with test validity. Transcript and diploma notations and denials that satisfy due process were presented as possibilities for test administrators and educators to consider. Phillips concluded by providing a list of guidelines for the review of all testing accommodation requests. The guidelines addressed request forms and procedures, documentation of disability, flagging of non-standard conditions, individualized decisions with evaluation criteria, administrative reviews and appeals, handbooks on the test accommodation process, and accommodation costs.

Phillips, S.E. (1994). High-stakes testing accommodations: Validity versus disabled rights. *Applied Measurement in Education, 7* (2), 93-120.

In this article, measurement problems associated with accommodations for learning disabilities and other high-prevalence disabilities in high stakes assessment were examined. Phillips constructed a legal framework for considering these accommodations by using existing case law, and she discussed the advantages and disadvantages of alternative strategies for handling testing accommodation requests, including self-selection with informed disclosure and the elimination of extraneous skills from assessments. Phillips suggested that there are three considerations to be made when determining whether to grant a requested accommodation: the purpose of the test, skills to be measured, and inferences to be made. She provided a list of five questions for measurement specialists to consider when determining whether to grant a requested accommodation and twelve recommendations for developing and implementing legally defensible testing accommodations policies.

Phillips, S.E. (1995). *All students, same test, same standards: What the new Title I legislation will mean for the educational assessment of special education students*. Oak Brook, IL: North Central Regional Educational Laboratory.

Prepared for the North Central Regional Educational Laboratory, this document is basically a shortened and condensed version of Philips' other three more comprehensive articles. In this report, Phillips considers what the impact of Title I legislation (which requires states to hold all students to the same expectations and to ensure they have equal educational opportunities) will be on state and local district evaluation plans. Phillips describes the tension between the seemingly competing goals of including all students with disabilities and the need to maintain educational standards in order for high school diplomas to be meaningful. She highlights two critical issues that need to be addressed: what kinds of allowances should be made and which

accommodations compromise technical adequacy (i.e., reliability and validity). The differences between physical and cognitive disabilities are briefly reviewed with respect to the issues surrounding accommodations made for each type of disability. Following this discussion Phillips reviews the legal precedent for accommodations including ADA, IDEA, Section 504, and relevant court cases. Phillips asserts that although none of the newer legislation has been tested yet in court, what *is* known is that in the past "judges have been deferential to academic decisions as long as procedural safeguards are followed. The courts have reinforced the quality issue; schools do not have to lower standards" (p. 5). She concludes by discussing the implications of this for policy formation.

Phillips, S.E. (1996). Legal defensibility of standards: Issues and policy perspectives. *Educational Measurement: Issues and Practice, 15* (2), 5-19.

In this article, Phillips describes legal criteria that may be challenged as states develop and implement high stakes assessments and graduation standards. Specifically, she defines and discusses six legal criteria for "descriptive standards" (i.e., goal statements describing what students should know and be able to do in specific content subjects): Notice, curricular, adverse impact, opportunity for success or fundamental fairness, articulating defensible standards, and assessment accommodations for students with disabilities.

Phillips describes legal precedent that has arisen from other case law, reviews professional requirements in these areas (e.g., Standards for Educational and Psychological Testing and The Code of Fair Testing Practices), speculates about possible litigation that may arise as states and school districts struggle to develop assessments and policy in this area, and suggests that the most important requirement when considering the rights of students with disabilities versus validity requirements is the "development of a comprehensive written policy outlining the procedures for requesting accommodations and detailing how decisions will be made regarding specific requests" (p. 12).

Teacher and Student Perceptions of Classroom Testing Adaptations and Modification

Gajria, M., Salend, S.J., & Hemrick, M.A. (1994). Teacher acceptability of testing modifications for mainstreamed students. *Learning Disabilities Research and Practice, 9* (4), 236-243.

Sixty-four general education teachers (grades 7-12) in two school districts in New York

completed a questionnaire focusing on awareness, use, integrity, effectiveness, and ease of use for 32 test design modifications. Results indicated that most teachers were familiar with many of the modifications and they were most likely to use modifications that could be applied to all students. They were less likely to use modifications that were specific to the needs of individual students. Modifications pertaining to changes in test design were more likely to be used than those requiring changes to administrative procedures. In addition, teachers were not likely to use modifications they believed endangered test integrity. For approximately one-third of the items, perceived effectiveness was rated significantly higher than use. The authors suggest that perceptions of effectiveness and the resources required for implementation are related to teachers' use of modifications. Discussion focused on the need for teacher training, education, and support around the incorporation of testing modifications for students with disabilities.

Jayanthi, M., Bursuck, W.D., Havekost, D.M., Epstein, M.H., & Polloway, E.A. (1994). School district testing policies and students with disabilities: A national survey. *School Psychology Review, 23* (4), 694-703.

As a first step toward understanding testing practices in general education classrooms, 214 school districts across the United States were surveyed regarding their testing policies. Questions focused on whether districts had formal policies concerning standardized and nonstandardized testing, what areas and issues the policies addressed, and what provisions were included in the policies for students with disabilities. Results indicated that nearly 60% of schools had formal policies for standardized tests. With respect to students with disabilities, 61% of the schools with formal policies reported that their policies require modifications for students with disabilities on standardized tests. Students with severe disabilities, followed by students with learning disabilities, were most often exempted from testing. The most frequently used modification was the use of a special administrator. Other modifications included use of large print, aides, a special site, and extended time. Forty-nine percent of the schools reported that they are required to report standardized test scores for students with disabilities. Very different results were obtained regarding nonstandardized testing policies. Only 22% of schools had formal policies, of those, 56% are required to make modifications for students with disabilities. Modifications most frequently used in these settings include a special administrator followed by extended time. Though small sample size (due to low response rate) limits the generalizability of conclusions, the authors conclude that while testing reform may be a national priority, it has yet to have a significant impact at the local level.

Jayanthi, M., Epstein, M.H., Polloway, E.A., & Bursuck, W.D. (1996). A national survey of general education teachers' perceptions of testing adaptations. *The Journal of Special Education, 30* (1), 99-115.

In this national study, 401 general education teachers (elementary through high school) were surveyed on their perceptions of testing adaptations for students with disabilities in general education classrooms. Results indicated that for the majority of respondents (83%), general educators, either alone or jointly with a special educator, were responsible for making decisions about testing adaptations in the classrooms. When asked to rate testing adaptations on scales indicating helpfulness to the student and ease of implementation, most of the adaptations rated as most helpful were not rated as easy to make. Examples of adaptations rated most helpful included "giving individual help with directions on a test" and "simplifying wording of test questions." "Allowing answers in outline formats" and "giving take-home tests" were rated as some of the least helpful adaptations. Items such as "using black-and-white copies instead of dittos" and "giving individual help with directions on a test" were rated as easy adaptations to make while "teaching students test-taking skills" and "allowing word processors" were rated among the most difficult to implement. The majority of teachers (67%) believed that it is unfair to provide testing adaptations only for students with identified disabilities; many of them stated that adaptations should be made for all students who need them. A small percentage of teachers (8%) believed that adaptations were unfair because all students in general education should work on general education standards. Differences also were found across grade levels. Consistently, teachers at the elementary level reported that adaptations were easier to implement. The authors conclude that discussions are limited by the fact that there is so little empirical research examining the effectiveness of testing adaptations for students with disabilities.

Schumm, J.S. & Vaughn, S. (1991). Making adaptations for mainstreamed students: General classroom teachers' perspectives. *Remedial and Special Education, 12* (4), 18-27.

Ninety-three elementary (N=25), middle school (N=23) and high school (N=45) teachers from one metropolitan school district in the Southeastern United States were asked to rate both the desirability and feasibility of 30 classroom adaptations using the Adaptation Evaluation Instrument (AEI). Results from both studies indicated that ratings of desirability were significantly higher than ratings of feasibility for all 30 adaptations. Adaptations that required little individualization were rated as most feasible to implement. Conversely, adaptations requiring changes in planning, curriculum use, or evaluation procedures were rated as least feasible. The adaptations rated most desirable by teachers were those that related to students' social and motivational adjustment and did not require any curricular or environmental

adaptations by the teacher (p. 22). Teachers in this study rated adaptations to materials or instruction as neither desirable nor feasible. There were very few differences as a function of grade. According to the authors, the bottom line for successful inclusion is "teacher willingness to accept and make decisions for students with special needs" (p. 18). The results of their study suggest that this may not be realistic given teachers' current level of information and skill in terms of making individualized adaptations, and they point to the need for ongoing teacher education, training, and support.

Vaughn, S., Schumm, J.S., Niarhos, F.J., & Daugherty, T. (1993). What do students think when teachers make adaptations? *Teaching and Teacher Education, 9* (1), 107-118.

In this study, middle and high school students (N=876) were surveyed on their perceptions of teacher-made adaptations. Results indicated that though students preferred teachers who made adaptations, they preferred certain types of adaptations. With respect to instructional practices, most students preferred teachers who were attentive to individual needs, sensitive to diverse learning patterns, and who adjusted instruction to meet the ability level of the student. Students preferred no adaptations in terms of textbooks, materials, homework, or tests. These results held when students were divided into high and low achieving groups and the authors found that high achieving students were not resentful of adaptations made for lower achieving students. The authors speculate that student preferences are related to the appearance of differential treatment. That is, students are less supportive of adaptations that "overtly indicate differential treatment" (p. 115) and this effect appears to intensify as students move from middle to high school. The researchers also surveyed students about achievement and social alienation in order to measure the relationship between these variables and students' perceptions of teacher adaptations. Results of these analyses showed that students who felt more alienated from their peers and teachers are more likely to hold favorable views of teachers who make adaptations. Unfortunately, the authors did not indicate whether their sample consisted of students in regular education, special education, or some combination of both; thus, generalizations about these findings must be made with caution.

Conceptual Issues Related to Testing and Accommodations

Bennett, R.E. (1995). *Computer-based testing for examinees with disabilities: On the road to generalized accommodation* (RM-95-1). Princeton, NJ: Educational Testing Service.

The use of computer-based testing (CBT) is proposed as one way to move toward generalized

testing accommodations, accommodations that can be used for all students, not just for students with disabilities. Bennett reviews the current practice of flagging tests taken through nonstandard administrations as well as research conducted by Willingham et al. (1988) indicating that the primary source of noncomparability for paper-and-pencil tests is the accommodation of extended time. He suggests that while task and score comparability have not yet been achieved for pencil-and-paper tests, CBT offers promise, particularly in terms of task comparability. Given that many people with disabilities already use computers as a life-style accommodation, using them for assessment purposes appears to be a natural extension that allows examinees to interact with the test in a variety of ways (e.g., using a head-mounted mouse emulator that allows examinees to magnify screens). Bennett also suggests that CBTs can possibly change timing constraints that have traditionally resulted from conducting group administrations of exams. As long as speededness is not an objective to be measured, the fact that CBTs are administered individually could, according to Bennett, resolve some of the current issues around timing. As yet, there is no definitive research on score comparability for CBTs.

Camara, W.J. & Brown, D.C. (1995). Educational and employment testing: Changing concepts in measurement and policy. *Educational Measurement, 14* (1), 5-11.

This article discusses current measurement, policy and social issues surrounding assessment in both educational and employment domains. Areas of educational testing and measurement that have experienced rapid and significant transformation and are relevant to employment testing are discussed, including expanding concepts of validity, the movement toward "authentic" or performance based assessment, and the changing uses and expectations of assessment and the implications for practice. The authors assert that there have been two important trends with respect to evolving views of validity. First, the centrality of construct validity has been established. Second, there is increased consideration given to social consequences when evaluating test validity. Despite the fact that performance-based assessments are used in both employment and educational settings, the research on performance-based assessment in education is minimal while research from the employment sector provides substantive evidence of reliability and validity. Explanations of these differences revolve around the purposes of the assessments, legislative and technical standards that have spurred the development of quality rating methods in employment settings, and the specificity with which employment assessments can measure skills that are directly relevant for the job. Discussion of the changing views of uses and expectations of assessment focuses on three levels: decision making, aiding instruction, and accountability. The authors conclude that because basic principles of measurement apply to all realms where tests are used (p. 10), greater interaction and

participation among professionals from different domains can only serve to increase peoples' understanding of the relevant issues.

Fischer, R.J. (1994). The Americans with Disabilities Act: Implications for measurement. *Educational Measurement: Issues and Practices, 13* (3), 17-26.

Fischer provides an overview of the evolution of the Americans with Disabilities Act (ADA). Although this article focuses primarily on assessment for employment purposes, many of the issues overlap with educational assessment. Fischer clearly explains the intent and purpose of the ADA and provides a succinct history of the legislation and practice leading up to its enactment. The definition of "reasonable accommodations" is discussed including reasons for and the implications of such a broad and vague definition on measurement issues. General measurement issues regarding appropriate testing accommodations are presented and several suggestions are made regarding ways to work within the ADA guidelines.

Geisinger, K.F. (1994). Psychometric issues in testing students with disabilities. *Applied Measurement in Education, 7* (2), 121-140.

Geisinger reviews psychometric constructs, particularly validity, that should be considered if students take non-standardized forms of a standardized test. He reviews empirical studies that have examined non-standard administrations from the perspective of criterion-related, content-related, and construct-related validation, with reliability analyses described as well. In addition, he discusses the constructs of fairness, differential item functioning, robustness, testing compensatory skills, and testing comparability. According to Geisinger, the greatest need is for more empirical research, both post-test validation studies and pre-test developmental research on the effects of various modifications.

Hishinuma, E.S. (1995). WISC-III accommodations: The need for practitioner guidelines. *Journal of Learning Disabilities, 28* (3), 130-135.

This article begins with a discussion of general issues surrounding testing accommodations, namely tensions between the competing goals of providing reasonable accommodations and the effects of nonstandardized testing procedures, as well as ADA requirements for testing. With these issues in mind, Hishinuma focuses on the WISC-III and eloquently discusses the pressing need for guidelines. Asserting that practitioners often have to make individual

decisions about accommodations on-the-spot (often during testing) with little information or guidance, Hishinuma states that guidelines are needed in three specific areas: initial selection of tests to administer, modifications in administration, and interpretation and reporting of results. He reviews the options currently available to practitioners and suggests possible ideas for future research on administration modifications. Stating that "legislative intent goes well beyond any preexisting research knowledge of the psychometric effects of accommodations" (p. 134), Hishinuma calls for research and subsequent guidelines that are based on a synthesis of information sources including student needs, professional ethics, psychometric theory, and empirical research.

Linn, R.L. (1994a). Evaluating the technical quality of proposed national examination systems. *American Journal of Education, 102* (4), 565-580.

This article examines the technical and measurement issues that arise when national examination systems are proposed. Although there is a lack of consensus about the purpose of national assessments, there is some clarity about measurement and technical quality issues associated with this type of assessment. Validity is identified as the most important technical concern and other technical quality issues (e.g., fairness and generalization) are considered to be important primarily because they have a direct bearing on validity. According to Linn, five types of evidence should be gathered in order to make informed judgments regarding the use and interpretation of results of a national examination system. These include information on content analysis, fairness, impact analysis, generalizability, and comparability. In all cases, the stringency of the validity evaluation is driven by the stakes of the assessment. The importance of representativeness with respect to statistical aggregates is noted and Linn concludes that the best bet for a national examination system (in terms of satisfying both policymakers and measurement specialists) will be to utilize a "phased-in system that begins with low stakes and increases those stakes only after the technical quality of the assessments has been adequately established" (p. 578).

Linn, R.L. (1994b). Performance assessment: Policy promises and technical measurement standards. *Educational Researcher, 23* (9), 4-14.

The expanded role of the federal government, the increased emphasis on standards, and the major increase in reliance on performance-based assessments are three of the changes in the nature and contexts of assessment described by Linn. A result of these changes is that efforts to revise and update the *Standards for Educational and Psychological Testing* (AERA, APA,

NCME, 1985) are now underway. Issues related to the certification of state performance assessments are presented including the role of standards and the political context in which decisions about standards and educational assessments are being made. Following this discussion, Linn shifts to technical measurement standards, specifically validity. He reviews the areas for which there is consensus in the measurement community (e.g., the primacy of validity, validity as a unitary construct) as well as discussing the areas that may be challenging for the revised standards to address (e.g., the consequences of assessments and the growing support for the inclusion of consequences as validity evidence). Linn asserts that as a result of Goals 2000 legislation, there are several emergent issues for the revised standards to address, including: performance and opportunity-to-learn standards, issues specific to students with IEPs or Limited English Proficiency (LEP), and reliability and generalization issues related to performance assessment. Linn emphasizes that the new standards need to be emphatic about the importance of validity so that it is not just a slogan. He argues that a better way must be found for establishing priorities for obtaining evidence about assessments (p. 13).

Messick, S. (1994, October). Alternative modes of assessment, uniform standards of validity. Research Report. Paper presented at Conference on Evaluating Alternatives to Traditional Testing for Selection, Bowling Green, Ohio. ERIC #ED380504.

Focusing on performance assessments as alternatives to multiple choice assessments, Messick posits that all forms of assessment should be held to a unitary concept of validity. Messick argues that inferences and action implications drawn from alternative assessments like performance assessments are fundamentally similar. The notion of a continuum is put forth to represent the range between multiple choice and performance assessments. The differences between construct and task-driven performance assessments are described and two major sources of invalidity – construct-irrelevant variance (authenticity) and construct under-representation (directness) – are presented. Claiming that validity questions commonly asked by measurement specialists generally seek to establish construct validity, Messick describes six distinguishable aspects of construct validity: content, substantive, structural, generalizability, external, and consequential. He concludes that a unitary concept of validity implies an integration of multiple supplementary forms of evidence, not answering just one question or providing evidence of one or more aspects of the assessment.

Overton, G.R. (1991, February). Accommodation of disabled persons. *The Bar Examiner*, 6-10.

This article, written before the enactment of the Americans with Disabilities Act, presents a framework for the implementation of a system of accommodations for persons with disabilities taking the bar exam. Key points cover:

- The inclusion of people with and without disabilities on committees to develop guidelines and plans for accommodations.
- Early notification to students about the availability of accommodations and early decisions about requested bar exam accommodations.
- Development of an appeal and review board comprised of persons with and without disabilities.
- Outreach to members of the bar for financial support for accommodations and education about accommodations and the need for them.

While not directly relevant to the discussion of high stakes graduation assessments, this article highlights the need to include persons with disabilities in planning and designing assessment systems. In addition, research conducted with previous bar examinees suggests that additional time may act as a stressor due to fatigue and stamina issues, an issue that may need consideration for graduation standards exams.

Pomplun, M. (1996). Cooperative groups: Alternative assessment for students with disabilities. *The Journal of Special Education, 30* (1), 1-17.

In this study, students with disabilities who participated in the Kansas Science Assessment were compared to students without disabilities. Given in fifth grade, the science assessment consists of both objective measures of performance and a cooperative group project. The purposes of the study were to determine whether assessment scores for students with disabilities were consistent with educators' expectations (e.g., students with moderate mental handicaps would score lower than students with speech difficulties) as well as determining whether the science scores measured the same abilities for students with and without disabilities. Another purpose was to examine whether the presence of a student with a disability affected group performance in terms of achievement. Results indicated that scores for students with disabilities were consistent with expectations; however the comparability of abilities measured was questionable. There was no evidence that participation in the cooperative groups

by students with disabilities negatively affected group scores but the authors did note concerns that students with disabilities may have been excluded from the group process (as evidenced by relatively low ratings of cooperation combined with higher than predicted group scores for those groups containing one or more students with a mental disability). Future research areas are suggested by the authors including the nature of student participation, the consequences of participation, and the appropriate use of resulting scores. One limitation to this study is that some accommodations were provided to students with disabilities in order to allow them to participate in the assessment process but they were not described.

Ragosta, M. (1991, February). Testing bar applicants with learning disabilities. *The Bar Examiner*, 11-15.

General policy issues related to testing students with disabilities as well as specific considerations for students with learning disabilities are reviewed in this article. In 1991, there were no validity data available on the effects of accommodations for the bar exam. Consequently, Ragosta spoke of the need for research and a more standardized policy for reviewing and granting accommodations. In terms of general guidelines, eligibility and documentation issues were discussed. Ragosta contends that requests for special accommodations should consider two issues: past educational practice and accommodation in the profession. With respect to candidates with learning disabilities, Ragosta reviews definitions of learning disabilities and discusses the significance of timing of diagnosis as a potentially important eligibility issue. Current accommodations for students with learning disabilities taking the SAT or LSAT are presented followed by suggestions for possible accommodations for people taking the bar exam, including:

- Alternate versions of the exam.
- Personal assistance for reading questions or recording answers.
- Assistive devices (e.g., tape recorder)
- A separate room.
- Extra time.

Given that extra time has been found to be the only accommodation thus far that produces noncomparable scores (Willingham et al., 1988), Ragosta also offers suggestions for setting time limits for candidates with learning disabilities based on level of disability and documentation of prior accommodations.

Salend, S.J. (1995). Modifying tests for diverse learners. *Interventions in Schools and Clinics, 31* (2), 84-90.

Recognizing that students with learning and behavioral disabilities may struggle with teacher-made assessments, this article presents a variety of modifications that teachers of students with disabilities can make to their tests in order to better meet their students' needs. Suggestions are made with respect to format changes, the presentation of items and directions, response options, two-tiered testing, short answer and essay items, and the readability of items. Salend also describes alternative grading systems, the need for collaboration between general and special educators and the limitations of teacher-made tests. Though this article is filled with suggestions for adaptations and modifications, there is very little discussion of issues related to technical adequacy other than a brief description of reliability and validity, as well as a cautionary note to teachers to carefully evaluate reliability and validity whenever they are designing or modifying tests. Although this article provides little to no guidance regarding modifications for high stakes assessments, it does provide a plethora of suggestions that may be helpful for thinking creatively about modifications that could be explored for higher stakes assessments.

Willingham, W.W. (1989). Standard testing conditions and standard score meaning for handicapped examinees. *Applied Measurement in Education, 2* (2), 97-103.

In this article, Willingham reviews the following guidelines on adaptations and modifications made to standardized tests: Section 504 regulations, the Joint Test Standards, The Panel on Testing of Handicapped People, and the psychometric literature. In addition, he defines and discusses score and task comparability, briefly describing each of the eight marks of comparability: factor structure, item functioning, reliability, predicted performance, admissions decisions, test content, testing accommodations, and test timing. Willingham then provides a concise summary of the research he and his colleagues conducted on the effects of nonstandard administrations of the SAT and GRE on students with disabilities over a four-year period. Generally, the author and his colleagues found that other than time limits, tests administered to examinees with disabilities were largely comparable to those used in the regular administrations. Willingham concludes that with respect to judging the comparability of the task, "timing is the critical issue." He then discusses several ways to improve nonstandard tests and suggests ways to conduct research to develop empirically-based time limits.

References

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Americans with Disabilities Act (A.D.A., 1990). Pub. L. No. 101-336, 42 U.S.C. 12101 *et seq*.

Bennett, R.E. (1995). *Computer based testing for examinees with disabilities: On the road to generalized accommodation* (RM-95-1). Princeton, NJ: Educational Testing Service.

Camara, W.J. & Brown, D.C. (1995). Educational and employment testing: Changing concepts in measurement and policy. *Educational Measurement, 14* (1), 5-11.

Dalton, B., Morocco, C.C., Tivnan, T., & Rawson, P. (1994). Effect of format on learning disabled and non-learning disabled students' performance on a hands-on science assessment. *International Journal of Educational Research, 21* (3), 229-316.

Debra P. v. Turlington, 474 F. Supp. 244 (M.D. Fla. 1979), *aff'd in part, rev'd in part*, 644 F.2d 397 (5th Cir. 1981); *on remand*, 564 F. Supp. 177 (M.D. Fla. 1983), *aff'd*, 730 F.2d 1405 (11th Cir. 1984).

Dunlap, G., Foster-Johnson, L., Clarke, S., Kern, L., & Childs, K.E. (1995). Modifying activities to produce functional outcomes: Effects on the problem behaviors of students with disabilities. *Journal of the Association for Persons with Severe Handicaps, 20* (4), 248-258.

Erickson, R., Thurlow, M., & Ysseldyke, J. (1996). *Neglected numerators, drifting denominators, and fractured fractions: Issues in determining participation rates for students with disabilities in statewide assessment programs*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Fischer, R.J. (1994). The Americans with Disabilities Act: Implications for measurement. *Educational Measurement: Issues and Practices, 13* (3), 17-26.

Gajria, M., Salend, S.J., & Hemrick, M.A. (1994). Teacher acceptability of testing modifications for mainstreamed students. *Learning Disabilities Research and Practice, 9* (4), 236-243.

Geisinger, K.F. (1994). Psychometric issues in testing students with disabilities. *Applied Measurement in Education, 7* (2), 121-140.

Hawaii State Dept. of Educ., 17 EHLR 360 (O.C.R. 1990).

Hishinuma, E.S. (1995). WISC-III accommodations: The need for practitioner guidelines. *Journal of Learning Disabilities, 28* (3), 130-135.

Jayanthi, M., Bursuck, W.D., Havekost, D.M., Epstein, M.H., & Polloway, E.A. (1994). School district testing policies and students with disabilities: A national survey. *School Psychology Review, 23* (4), 694-703.

Jayanthi, M., Epstein, M.H., Polloway, E.A., & Bursuck, W.D. (1996). A national survey of general education teachers' perceptions of testing adaptations. *The Journal of Special Education, 30* (1), 99-115.

Joint Committee on Testing Practices. (1988). *Code of fair testing practices in education*. Washington, DC: Author.

Linn, R.L. (1994a). Evaluating the technical quality of proposed national examination systems. *American Journal of Education, 102* (4), 565-580.

Linn, R.L. (1994b). Performance assessment: Policy promises and technical measurement standards. *Educational Researcher, 23* (9), 4-14.

Maxwell v. Pasadena I.S.D., No. 92-017184, 295th District Court of Harris County, TX, Dec. 29, 1994.

Messick, S. (1994). *Alternative modes of assessment, uniform standards of validity*. Princeton, N.J.: Educational Testing Service (ED 380 504).

Mick, L.B. (1989). Measurement effects of modifications in minimum competency test formats for exceptional students. *Measurement and Evaluation in Counseling and Development, 22*, 31-36.

Munger, G.F. & Loyd, B.H. (1991). Effect of speededness on test performance of handicapped and nonhandicapped examinees. *Journal of Educational Research, 85* (1), 53-57.

Overton, G.R. (1991, February). Accommodation of disabled persons. *The Bar Examiner*, 6-10.

Perlman, C., Borger, J., Collins, C., Elenbogen, J., & Wood, J. (1996, April). *The effect of extended time limits on learning disabled students' scores on standardized reading tests*. Paper presented at the annual meeting of the National Council on Measurement in Education, New York, New York.

Phillips, S.E. (1993). Testing accommodations for disabled students. *Education Law Reporter, 80*, 9-32.

Phillips, S.E. (1994). High-stakes testing accommodations: Validity versus disabled rights. *Applied Measurement in Education, 7* (2), 93-120.

Phillips, S.E. (1995). *All students, same test, same standards: What the new Title I legislation will mean for the educational assessment of special education students*. Oak Brook, IL: North Central Regional Educational Laboratory.

Phillips, S.E. (1996). Legal defensibility of standards: Issues and policy perspectives. *Educational Measurement: Issues and Practice, 15* (2), 5-19.

Pomplun, M. (1996). Cooperative groups: Alternative assessment for students with disabilities. *The Journal of Special Education, 30* (1), 1-17.

Ragosta, M. (1980). *Handicapped students and the SAT* (Research Report 80-12). Princeton, NJ: Educational Testing Service.

Ragosta, M. (1991). Testing bar applicants with learning disabilities. *The Bar Examiner, February*, 11-15.

Ragosta, M. & Wendler, C. (1992). *Eligibility issues and comparable time limits for disabled and nondisabled SAT examinees*. New York, N.Y.: College Entrance Examination Board Report No. 92-5. (ED 349 337).

Schumm, J.S. & Vaughn, S. (1991). Making adaptations for mainstreamed students: General classroom teachers' perspectives. *Remedial and Special Education, 12* (4), 18-27.

Salend, S.J. (1995). Modifying tests for diverse learners. *Interventions in Schools and Clinics, 31* (2), 84-90.

Section 504 of the Rehabilitation Act, 29 U.S.C. 701 *et seq*. (1973).

Thurlow, M.L., Ysseldyke, J.E., & Silverstein, B. (1993). *Testing accommodations for students with disabilities: A review of the literature* (Synthesis Report 4). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Vaughn, S., Schumm, J.S., Niarhos, F.J., & Daugherty, T. (1993). What do students think when teachers make adaptations? *Teaching and Teacher Education, 9* (1), 107-118.

Wards Grove Packing Co. v. Antonio, 109 S.Ct. 2115 (1989).

Willingham, W.W. (1989). Standard testing conditions and standard score meaning for handicapped examinees. *Applied Measurement in Education, 2* (2), 97-103.

Appendix A

Research Projects Supported by the U.S. Office of Special Education Programs (OSEP) and the U.S. Office of Educational Research and Improvement (OERI)

Research Projects Supported by the U.S. Office of Special Education Programs (OSEP) and the U.S. Office of Educational Research and Improvement (OERI)

OSEP Projects	*Recipient Organizations*
Examining Alternatives for Outcome Assessment for Children with Disabilities	Maryland State Department of Education
Performance Assessment and Standardized Testing for Students with Disabilities: Psychometric Issues, Accommodation Procedures, and Outcome Analyses	Wisconsin Center for Education Research
Project Reading ABC: An Alternative Reading Assessment Battery for Children with Severe Speech and Physical Impairments	Center for Literacy and Disability Studies - University of North Carolina

OERI Projects	*Recipient Organizations*
Inclusive Comprehensive Assessment System	Delaware Department of Public Education
The Maryland Assessment System Project	Maryland State Department of Education
Grade 5 and 8 Integrated Social Studies Statewide Assessment Project	Michigan Department of Education
Minnesota Assessment Project	Minnesota Department of Education
Assessment of Media Literacy	North Carolina Department of Public Instruction
North Dakota Language Arts Assessment	North Dakota Department of Public Instruction
Oregon Assessment Development and Evaluation Project	Oregon Department of Education
Pennsylvania Assessment Through Themes Project	Pennsylvania Department of Education
State Collaborative on Assessment and Student Standards (SCASS) Technical Guidelines for Performance Assessment	Council of Chief State School Officers (CCSSO)

Appendix B

References From the Report, *Testing Accommodations for Students with Disabilities: A Review of the Literature* (Thurlow, Ysseldyke, & Silverstein, 1993)

References

American Psyhchological Association (1985). *Standards for educational and psychological testing*. Washington, DC: Author.

Amos, K. M. (1980). Competency testing: Will the LD student be included? *Exceptional Children, 47* (3), 194-197

Beattie, S., Grise, P., & Algozzine, B. (1983). Effects of test modifications on the minimum competency performance of learning disabled students. *Learning Disability Quarterly, 6* (1), 75-77.

Benderson, A. (Ed.) (1988) Testing, equality, and handicapped people. *Focus* (ETS Publication; Princeton, NJ), *21* (23 pp).

Bennett, R. E., Rock, D. H., & Jirele, T. (1987). GRE score level, test completion, and reliability for visually impaired, physically handicapped, and nonhandicapped groups. *The Journal of Special Education, 21* (3), 9-21.

Bennett, R. E., Rock, D. A., & Jirele, T. (1986, February). *The psychometric characteristics of the GRE General Test for three handicapped groups* (ETS Research Report RR-86-6). Princeton, NJ: Educational Testing Service.

Bennett, R. E., Rock, D. A., & Kaplan, B. A. (1985, November). *The psychometric characteristics of the SAT for nine handicapped groups* (ETS Research Report RR-85-49). Princeton, NJ: Educational Testing Service.

Bennett, R. E., Rock, D. A., & Kaplan, B. A. (1987). SAT differential item performance for nine handicapped groups. *Journal of Educational Measurement, 24* (1), 44-55.

Bennett, R. E., Rock, D. A., & Kaplan, B. A. (1988). Level reliability and speededness of SAT scores for nine handicapped groups. *Special Services in the Schools, 4* (3/4), 37-54.

Bennett, R. E., Rock, D. A., Kaplan, B. A., & Jirele (1988). Psychometric characteristics. In W. W. Willingham, M. Ragosta, R. E. Bennet, H. Braun, D. A. Rock, & D. E. Powers (Eds.) *Testing handicapped people* (pp. 83-97). Boston: Allyn & Bacon.

Braun, H., Ragosta, M., & Kaplan, B. (1986, November). *The predictive validity of the GRE General Test for disabled students* (ETS Research Report 86-42). Princeton, NJ: Educational Testing Service.

Braun, H., Ragosta, M., & Kaplan, B. (1986, October). *The predictive validity of the scholastic aptitude test for disabled students* (ETS Research Report RR-86-38). Princeton, NJ: College Entrance Examination Board, Educational Testing Service, Graduate Record Examinations Board.

Casey, E. (1987). Accommodating testing to disabled students. In D. Bray & M. J. Belcher (Eds.). New directions for community colleges, No. 59. *Issues in Student Assessment*. San Francisco: Jossey-Bass.

Danielson, L. C. (1980). Educational goals and competency testing for the handicapped. In R. M. Jaeger & C. K. Tittle (Eds.), *Minimum competency achievement testing: Motives, models measures, and consequences* (pp. 201-204). Berkeley, CA: McCutchan.

DeStefano, L., & Metzer, D. (1991). High stakes testing and students with handicaps: An analysis of issues and policies. In R. E. Stake (Ed.), *Advances in program evaluation* (Vol. 1A). Greenwich, CT:: JAI Press.

Division of Evaluation, Measurement, and Statistics. (1992). *Statement of the American Psychological Association Division of Evaluation, Measurement, and Statistics on psychometric and assessment issues raised by the Americans with Disabilities Act (ADA).* Unpublished draft.

Driscoll, N. B. (1985). Minimum competency testing of the handicapped: A review of legal challenges. *Special Services in the Schools, 2* (1), 95-101.

ETS. (1990). Testing persons with disabilities: A report for ETS programs and their constituents. Princeton, NJ: Educational Testing Service.

ETS. (1991). *NTE programs: 1991-92 bulletin of information*. Princeton, NJ: Educational Testing Service.

ETS. (1992). ETS conference examines the technology of computer-based testing for people with disabilities. *ETS Developments*.

Ewing, N. J., & Smith, J. E. (1981). Minimum competency testing and the handicapped. *Exceptional Children, 47* (7), 523-524.

FairTest. (1993). *Computerized testing: More questions than answers*. Cambridge, MA: National Center for Fair & Open Testing.

Fenton, K. S. (1980). Competency testing and the handicapped: Some legal concerns for school administrators. In R. M. Jaeger & C. K. Tittle (Eds.), *Minimum competency achievement testing: Motives, models, measures, and consequences* (pp. 182-188). Berkeley, CA: McCutchan.

Grise, P. J. (1980). Florida's minimum competency testing program for handicapped students. *Exceptional Children, 47* (3), 186-191.

Haladyna, T. M., Nolen, S. B., & Haas, N. S. (1991). Raising standardized achievement test scores and the origins of test score pollution. *Educational Researcher, 20* (5), 2-7.

King, W. L., & Jarrow, J. E. (1992). *Testing accommodations for persons with disabilities: A guide for licensure, certification, and credentialing*. Columbus, OH: Association on Higher Education and Disability (AHEAD).

Laing, J., & Farmer, M. (1984). *Use of the ACT assessment by examinees with disabilities* (Research Report No. 84). Iowa City, IA: American College Testing Program (ACT).

Maxey, E. J., & Levitz, R. S. (1980, April). *ACT services for the handicapped*. Paper presented at the meeting of the American Association of Collegiate Registrars and Admissions Officers, New Orleans, LA.

McCarthy, M. M. (1980). Minimum competency testing and handicapped students. *Exceptional Children, 47* (3), 166-173.

McGrew, K. S., Thurlow, M. L., Shriner, J. G., & Spiegel, A. N. (1992). *Inclusion of students with disabilities in national and state data collection programs* (Technical Report 2). Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

Morrisey, P. (1980). Adaptive testing: How and when should handicapped students be accommodated in competency testing programs? In R. M.. Jaeger C. K. title (Eds.), Minimum competency achievement testing: Motives, models, and consequences. (pp. 205-210). Berkeley, CA: McCutchan

NASDSE. (1988). *NAEP testing for state comparisons: Issues related to inclusion of handicapped students (Executive Summary)*. Washington, DC: Author.

Office of Technology Assessment. (1987). *State educational testing practices*. (Unpublished background paper of the Science, Education, and Transportation Program). Washington, DC: Author.

Office of Technology Assessment. (1992). *Testing in American schools: Asking the right questions*. Washington, DC: Author.

Olsen, K. (1980). Minimum competency testing and the IEP process. *Exceptional Children, 47* (3), 176-183.

Pelavin Associates, Inc. (1991). *Statewide testing programs*. Unpublished document presented to the Secretary's Commission on Achieving Necessary Skills. (Address: 2030 M Street, NW, Washington, DC 20036)

Phillips, S. E. (1992). *Testing condition accommodations for handicapped students*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Rock, Bennett, & Jirele (1988). Factor structure of the Graduate Record Examinations general test in handicapped and nonhandicapped groups. *Journal of Applied Psychology, 73* (3), 383-392.

Rock, Bennett, & Kaplan (1987). Internal construct validity of a college admissions test across handicapped and nonhandicapped groups. *Educational and Psychological Measurement, 47*, 193-205.

Safer, N. (1980). Implications of minimum competency standards and testing for handicapped students. *Exceptional Children, 46*, 288-290.

Serow, R. C., & O'Brien, K. O. (1983). Performance of handicapped students in a competency testing program. *The Journal of Special Education, 17* (2), 149-155.

Smith, J. D., & Jenkins, D. S. (1980). Minimum competency testing and handicapped students. *Exceptional Children, 46* (6), 440-443.

Suran, B. G., & Rizzo, J. V. (1983). *Special children: An integrative approach*. Glenview, Illinois: Scott, Foresman & Co.

Thurlow, M. L., & Ysseldyke, J. E. (1992, November). *Can "all" ever really mean "all" in defining and assessing student outcomes?* Paper presented at the annual meeting of the American Evaluation Association, Seattle.

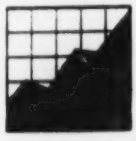
Vocational Rehabilitation Act, Public Law 93-112, Section 504, 1973, U.S. Congress.

Wildemuth, B. M. (1983). Minimum competency testing and the handicapped. Princeton, NJ: ERIC Clearinghouse on Tests, Measurement, and Evaluation.

Willingham, W. W. (1988). Introduction. In W. W. Willingham, M. Ragosta, R. E. Bennett, H. Braun, D. A. Rock, & D. E. Powers (Eds.), *Testing handicapped people*. Boston: Allyn & Bacon.

Willingham, W. W., Ragosta, M, Bennett, R. E., Braun, H., Rock, D. A., & Powers, D. E. (Eds.). (1988). *Testing handicapped people*. Boston: Allyn & Bacon.

Wolman, C., Bruininks, R. H., & Thurlow, M. L. (1989). Dropouts and dropout programs: Implications for special education. *Remedial and Special Education, 10* (5), 6-20.



NATIONAL CENTER ON EDUCATIONAL OUTCOMES

University of Minnesota • 350 Elliott Hall
75 East River Road • Minneapolis, MN 55455
612/626-1530 • Fax 612/624-0879
http://www.coled.umn.edu/NCEO

The College of Education & Human Development
UNIVERSITY OF MINNESOTA

BEST COPY AVAILABLE